

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International patent Classification 4: C07K 13/00, C12N 15/00, C12P 21/02 A61L 27/00, A61K 37/36 // A61K 35/32		A2	(11) International Publication Number: WO 89/09788 (43) International Publication Date: 19 October 1989 (19.10.89)
(21) International Application Number: PCT/US89/01469 (22) International Filing Date: 7 April 1989 (07.04.89) (30) Priority data: 179,406 8 April 1988 (08.04.88) US 232,630 15 August 1988 (15.08.88) US 315,342 23 February 1989 (23.02.89) US (60) Parent Applications or Grants (63) Related by Continuation US 179,406 (CIP) Filed on 8 April 1988 (08.04.88) US 232,630 (CIP) Filed on 15 August 1988 (15.08.88) US 315,342 (CIP) Filed on 23 February 1988 (23.02.88) (71) Applicant (for all designated States except US): CREATIVE BIOMOLECULES, INC. [US/US]; 35 South Street, Hopkinton, MA 01748 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): OPPERMAN, Hermann [US/US]; 25 Summer Hill Road, Medway, MA 02053 (US). KUBERASAMPATH, Thangavel [IN/US]; 6 Spring Street, Medway, MA 02053 (US). RUEGER, David, C. [US/US];			150 Edgemere Road, Apt. 4, West Roxbury, MA 02132 (US). OZKAYNAK, Engin [TR/US]; 44 Purdue Drive, Milford, MA 01757 (US). (74) Agent: PITCHER, Edmund, R.; Lahive & Cockfield, 60 State Street, Boston, MA 02109 (US). (81) Designated States: AT (European patent), AU, BB, BE (European patent), BF (OAPI patent), BG, BJ (OAPI patent), BR, CF (OAPI patent), CG (OAPI patent), CH (European patent), CM (OAPI patent), DE (European patent), DK, FI, FR (European patent), GA (OAPI patent), GB (European patent), HU, IT (European patent), JP, KP, KR, LK, LU (European patent), MC, MG, ML (OAPI patent), MR (OAPI patent), MW, NL (European patent), NO, RO, SD, SE (European patent), SN (OAPI patent), SU, TD (OAPI patent), TG (OAPI patent), US. Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: BIOSYNTHETIC OSTEOGENIC PROTEINS AND OSTEOGENIC DEVICES CONTAINING THEM			
(57) Abstract Disclosed are 1) osteogenic devices comprising a matrix containing osteogenic protein and methods of inducing endochondral bone growth and cartilage in mammals using the devices; 2) amino acid sequence data, amino acid composition, solubility properties, structural features, homologies and various other data characterizing osteogenic proteins, 3) methods of producing osteogenic proteins using recombinant DNA technology, and 4) osteogenically and chondrogenically active synthetic protein constructs.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	ML	Mali
AU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands
BG	Bulgaria	IT	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SN	Senegal
CH	Switzerland	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
DE	Germany, Federal Republic of	LU	Luxembourg	TG	Togo
DK	Denmark	MC	Monaco	US	United States of America
FI	Finland	MG	Madagascar		

BIOSYNTHETIC OSTEOGENIC PROTEINS AND
OSTEOGENIC DEVICES CONTAINING THEM

This invention relates to osteogenic devices, to synthetic genes encoding proteins which can induce osteogenesis in mammals and methods for their production using recombinant DNA techniques, to synthetic forms of osteogenic protein, and to bone and cartilage repair procedures using osteogenic device comprising the synthetic proteins.

Mammalian bone tissue is known to contain one or more proteinaceous materials, presumably active during growth and natural bone healing, which can induce a developmental cascade of cellular events resulting in endochondral bone formation. This active factor (or factors) has variously been referred to in the literature as bone morphogenetic or morphogenic protein, bone inductive protein, osteogenic protein, osteogenin, or osteoinductive protein.

The developmental cascade of bone differentiation consists of recruitment of mesenchymal cells, proliferation of progenitor cells, calcification of cartilage, vascular invasion, bone formation, remodeling, and finally marrow differentiation (Reddi (1981) Collagen Rel. Res. 1:209-226).

Though the precise mechanisms underlying these phenotypic transformations are unclear, it has been shown that the natural endochondral bone differentiation activity of bone matrix can be dissociatively extracted and reconstituted with inactive residual collagenous matrix to restore full bone induction activity (Sampath and Reddi, (1981) Proc. Natl. Acad. Sci. USA 78:7599-7603). This provides an experimental method for assaying protein extracts for their ability to induce endochondral bone in vivo.

This putative bone inductive protein has been shown to have a molecular mass of less than 50 kilodaltons (kD). Several species of mammals produce closely related protein as demonstrated by cross species implant experiments (Sampath and Reddi (1983) Proc. Natl. Acad. Sci. USA 80:6591-6595).

The potential utility of these proteins has been widely recognized. It is contemplated that the availability of the pure protein would revolutionize orthopedic medicine, certain types of plastic surgery, and various periodontal and craniofacial reconstructive procedures.

The observed properties of these protein fractions have induced an intense research effort in various laboratories directed to isolating and identifying the pure factor or factors responsible for osteogenic activity. The current state of the art of purification of osteogenic protein from

mammalian bone is disclosed by Sampath et al. (Proc. Natl. Acad. Sci. USA (1987) 80). Urist et al. (Proc. Soc. Exp. Biol. Med. (1984) 173:194-199) disclose a human osteogenic protein fraction which was extracted from demineralized cortical bone by means of a calcium chloride-urea inorganic-organic solvent mixture, and retrieved by differential precipitation in guanidine-hydrochloride and preparative gel electrophoresis. The authors report that the protein fraction has an amino acid composition of an acidic polypeptide and a molecular weight in a range of 17-18 kD.

Urist et al. (Proc. Natl. Acad. Sci. USA (1984) 81:371-375) disclose a bovine bone morphogenetic protein extract having the properties of an acidic polypeptide and a molecular weight of approximately 18 kD. The authors reported that the protein was present in a fraction separated by hydroxyapatite chromatography, and that it induced bone formation in mouse hindquarter muscle and bone regeneration in trephine defects in rat and dog skulls. Their method of obtaining the extract from bone results in ill-defined and impure preparations.

European Patent Application Serial No. 148,155, published October 7, 1985, purports to disclose osteogenic proteins derived from bovine, porcine, and human origin. One of the proteins, designated by the inventors as a P3 protein having a molecular weight of 22-24 kD, is said to have been purified to an essentially homogeneous state. This material is reported to induce bone formation when implanted into animals.

International Application No. PCT/087/01537, published January 14, 1988, discloses an impure fraction from bovine bone which has bone induction qualities. The named applicants also disclose putative bone inductive factors produced by recombinant DNA techniques. Four DNA sequences were retrieved from human or bovine genomic or cDNA libraries and apparently expressed in recombinant host cells. While the applicants stated that the expressed proteins may be bone morphogenic proteins, bone induction was not demonstrated. See also Urist et al., EP 0,212,474 entitled Bone Morphogenic Agents.

Wang et al. (Proc. Nat. Acad. Sci. USA (1988) 85: 9484-9488) discloses the purification of a bovine bone morphogenetic protein from guanidine extracts of demineralized bone having cartilage and bone formation activity as a basic protein corresponding to a molecular weight of 30 kD determined from gel elution. Purification of the protein yielded proteins of 30, 18 and 16 kD which, upon separation, were inactive. In view of this result, the authors acknowledged that the exact identity of the active material had not been determined.

Wozney et al. (Science (1988) 242: 1528-1534) discloses the isolation of full-length cDNA's encoding the human equivalents of three

polypeptides originally purified from bovine bone. The authors report that each of the three recombinantly expressed human proteins are independently or in combination capable of inducing cartilage formation. No evidence of bone formation is reported.

It is an object of this invention to provide osteogenic devices comprising matrices containing dispersed osteogenic protein capable of bone induction in allogenic and xenogenic implants. Another object is to provide synthetic osteogenic proteins capable of inducing endochondral bone formation in mammals, including humans. Yet another object is to provide genes encoding non-native osteogenic proteins and methods for their production using recombinant DNA techniques. Another object is to provide novel biosynthetic forms of osteogenic proteins and a structural design for novel, functional osteogenic proteins. Another object is to provide methods for inducing cartilage formation.

These and other objects and features of the invention will be apparent from the description, drawings, and claims which follow.

Summary of the Invention

This invention involves osteogenic devices which, when implanted in a mammalian body, can induce at the locus of the implant the full developmental cascade of endochondral bone formation and bone marrow differentiation. Suitably modified as disclosed herein, the devices also may be used to induce cartilage formation. The devices comprise a carrier material, referred to herein as a matrix, having the characteristics disclosed below, containing dispersed osteogenic protein in the form of a biosynthetic construct.

Key to these developments was the successful preparation of substantially pure osteogenic protein by purification from bone, the elucidation of amino acid sequence and structure data of the native osteogenic protein, and insights involving study of the DNA and amino acid sequences of the natural source product. A protocol was developed which results in retrieval of active, substantially pure osteogenic protein from mammalian bone. Investigation of the properties and structure of the native form osteogenic protein then permitted the inventors to develop a rational design for non-native forms, i.e., forms never before known in nature, capable of inducing bone formation. As far as applicants are aware, the constructs disclosed herein constitute the first instance of the design of a functional, active protein without preexisting knowledge of the active region of a native form nucleotide or amino acid sequence.

A series of consensus DNA sequences were designed with the goal of producing an active osteogenic protein. The sequences were based on partial amino acid sequence data obtained from the naturally sourced product and on observed homologies with unrelated genes reported in the literature, or the sequences they encode, having a presumed or demonstrated developmental function. Several of the biosynthetic consensus sequences have been expressed as fusion proteins in procaryotes, purified, cleaved, refolded, combined with a matrix, implanted in an established animal model, and shown to have endochondral bone-inducing activity. The currently preferred active proteins comprise sequences designated COP5, COP7, COP16, and OP1. The amino acid sequences of these proteins are set forth below.

```

COP5      1      10      20      30      40
           LYVDFS-DVGWDDWIVAPPGYQAFYCHGECPFPLAD
           50      60      70
           HFNSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA
           80      90      100
           ISMLYLDENEKVVLKNYQEMVVEGCGCR

COP7      1      10      20      30      40
           LYVDFS-DVGWNDWIVAPPGYHAFYCHGECPFPLAD
           50      60      70
           HLNSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA
           80      90      100
           ISMLYLDENEKVVLKNYQEMVVEGCGCR

COP16      1      10      20      30      40
           CRRHSLYVDFS-DVGWNDWIVAPPGYQAFYCHGECPFPLAD
           50      60      70
           HFNSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA
           80      90      100
           ISMLYLDENEKVVLKNYQEMVVEGCGCR
           -10
           PKHHSQRARKKNKN

```

- 8 -

```

                                -5
                                HQRQA
OP1      1      10      20      30      40
          CKKHELYVSFR-DLGWQDWIIAPEGYAAYYCEGECAFFLNS
                        50      60      70
          YMNATN--H-AIVQTLVHFINPET-VPKPCCAPTQLNA
                        80      90      100
          ISVLYFDDSSNVILKKYRNMVVRACGCH

```

In these sequences and all other amino acid sequences disclosed herein, the dashes (-) are used as fillers only to line up comparable sequences in related proteins, and have no other function. Thus, amino acids 45-50 of COP7, for example, are NHAVV. Also, the numbering of amino acids is selected solely for purposes of facilitating comparisons between sequences. Thus, for example, the DF residues numbered at 9 and 10 of COP5 and COP7 may comprise residues, e.g., 35 and 36, of an osteogenic protein embodying the invention. Various leader or trailer sequences may be attached to the operative active region provided the osteogenic or chondrogenic activity of the protein is not destroyed.

Thus, in one aspect, the invention comprises a protein comprising an amino acid sequence sufficiently duplicative of the sequence of COP5, COP7, COP16, or OP1 such that it is capable of inducing endochondral bone formation when properly folded and implanted in a mammal in association with a matrix. Some of these sequences induce cartilage,

but not bone. Also, the bone forming materials may be used to produce cartilage if implanted in an avascular locus, or if an inhibitor to full bone development is implanted together with the active protein. Thus, in another aspect, the invention comprises a protein less than about 200 amino acids long (for each chain) including a sequence sufficiently duplicative of the sequence of COP5, COP7, COP16, or OPl such that it is capable at least of cartilage formation when properly folded and implanted in a mammal in association with a matrix. The phrase "sufficiently duplicative", as used herein, is used to describe proteins having a degree of homology with the specific sequences disclosed herein and other, different amino acids but which nevertheless exhibit osteogenic or chondrogenic activity.

In one preferred aspect, these proteins comprise species of the generic amino acid sequences:

```

1      10      20      30      40      50
  LXVXFDXGWXXWXXXPXGXXAXYCXGXCPXXXXXXNHAAX
      60      70      80      90     100
QXXVXXNXXXPXXCCPXXXXXXXXLXXXXXXXXVXLXXYXXMXVXXCXCX

```

or

```

1      10      20      30      40      50
CXXXLXVXFDXGWXXWXXXPXGXXAXYCXGXCPXXXXXXNHAAX
      60      70      80      90     100
QXXVXXNXXXPXXCCPXXXXXXXXLXXXXXXXXVXLXXYXXMXVXXCXCX

```

where the letters indicate the amino acid residues of standard single letter code, and the Xs each represent any one of the 22 naturally occurring amino acid residues. Preferred amino acid sequences within the foregoing generic sequences are:

```

1      10      20      30      40      50
  LYVDFRDVGWNDWIVAPPGYHAFYCHGECPPFLADHLNSTNHAIV
    K S S L  QE VIS E FD Y  E A AY MPESMKAS  VI
    F E K I  DN   L   N S   Q  ITK F P   TL
      A     S     K
      60      70      80      90      100
QTLVNSVNP GKIPKACCVPTELSAISMLYLDENENVVLKKNYQDMVVEGCGCR
SI HAI SEQV EP  A  EQMNSLAI FFNDQDK I RK EE T DA H H
  RF   T   S      K DPV V  Y N S      H RN   RS
    N   S                      K       P   E

```

and

```

1      10      20      30      40      50
CKRHPLYVDFRDVGWNDWIVAPPGYHAFYCHGECPPFLADHLNSTNHAIV
RRRS K S S L  QE VIS E FD Y  E A AY MPESMKAS  VI
  KE F E K I  DN   L   N S   Q  ITK F P   TL
    Q     A     S     K
      60      70      80      90      100
QTLVNSVNP GKIPKACCVPTELSAISMLYLDENENVVLKKNYQDMVVEGCGCR
SI HAI SEQV EP  A  EQMNSLAI FFNDQDK I RK EE T DA H H
  RF   T   S      K DPV V  Y N S      H RN   RS
    N   S                      K       P   E

```

- 11 -

wherein each of the amino acids arranged vertically at each position in the sequence may be used alternatively in various combinations. Note that these generic sequences have 6 and preferably 7 cysteine residues where inter- or intramolecular disulfide bonds can form, and contain other critical amino acids which influence the tertiary structure of the proteins. These generic structural features are found in previously published sequences, none of which have been described as capable of osteogenic activity, and most of which never have been linked with such activity.

Particular useful sequences include:

Vg1	1	10	20	30	40
	CKKRHLVVEFK-DVGWQNWVIAPQGYMANCYGECPYPLTE				
		50	60	70	
	ILNGSN--H-AILQTLVHSIEPED-IPLPCCVPTKMSP				
	80	90	100		
	ISMLFYDNNDNVVLRHYENMAVDECGCR				
DPP	1	10	20	30	40
	CRRHSLYVDFS-DVGWDDWIVAPLGDAYYCHGKCPFPLAD				
		50	60	70	
	HFNSTN--H-AVVQTLVNNNNPGK-VPKACCVPTQLDS				
	80	90	100		
	VAMLYLNDQSTVVVLKNYQEMTVVGCGR				
CBMP-2a	1	10	20	30	40
	CKRHPLYVDFS-DVGWNDWIVAPPGYHAFYCHGECPPFLAD				
		50	60	70	
	HLNSTN--H-AIVQTLVNSVNS-K-IPKACCVPTELSA				
	80	90	100		
	ISMLYLDENEKVVLKNYQDMVVEGCGCR				

- 12 -

CBMP-2b 1 10 20 30 40
 CRRHSLYVDFS-DVGWNDWIVAPPGYQAFYCHGDCPFPLAD
 50 60 70
 HLNSTN--H-AIVQTLVNSVNS-S-IPKACCVPTLSA
 80 90 100
 ISMLYLDEYDKVVLKNYQEMVVEGCGCR

CBMP-3 1 10 20 30 40
 CARRYLKVDFA-DIGWSEWIIISPKSFDAYYCSGACQFPMPK
 50 60 70
 SLKPSN--H-ATIQSIVRAVGVPPIPEPCCVPEKMSS
 80 90 100
 LSILFFDENKNVVLKVYPNMTVESACR

COP1 1 10 20 30 40
 LYVDFQRDVGWDDWIIAPVDFDAYYCSGACQFPSAD
 50 60 70
 HFNSTN--H-AVVQTLVNNMNPBK-VPKPCCVPTLSA
 80 90 100
 ISMLYLDENSTVVLKNYQEMTVVGCGCR

COP3 1 10 20 30 40
 LYVDFQRDVGWDDWIVAPPGYQAFYCSGACQFPSAD
 50 60 70
 HFNSTN--H-AVVQTLVNNMNPBK-VPKPCCVPTLSA
 80 90 100
 ISMLYLDENEKVVLKNYQEMVVEGCGCR

COP4 1 10 20 30 40
 LYVDFS-DVGWDDWIVAPPGYQAFYCSGACQFPSAD
 50 60 70
 HFNSTN--H-AVVQTLVNNMNPBK-VPKPCCVPTLSA
 80 90 100
 ISMLYLDENEKVVLKNYQEMVVEGCGCR

Vgl is a known Xenopus sequence heretofore not associated with bone formation. DPP is an amino acid sequence encoded by a Drosophila gene responsible for development of the dorsoventral pattern. OPl is a region of a natural sequence encoded by exons of a genomic DNA sequence retrieved by applicants. The CBMPs are amino acid sequences comprising subparts of mammalian proteins encoded by genomic DNAs and cDNAs retrieved by applicants. The COPs are totally biosynthetic protein sequences expressed by novel consensus gene constructs, designed using the criteria set forth herein, and not yet found in nature.

These proteins are believed to be dimers. They appear not to be active when reduced. Various combinations of species of the proteins may exist as heterodimers or homodimers. As far as applicants are aware, the COP5, COP7, COP16, and OPl constructs constitute the first instances of the design of a bioactive protein without preexisting knowledge of the active region of a native form nucleotide or amino acid sequence.

The invention thus provides synthetic osteogenic protein produced using recombinant DNA techniques. The protein may include forms having varying glycosylation patterns, varying N-termini, a family of related proteins having regions of amino acid sequence homology, and active truncated or mutated forms of native protein, no matter how

- 14 -

derived. In view of this disclosure, skilled genetic engineers can design and synthesize genes which encode appropriate amino acid sequences, and then can express them in various types of host cells, including both procaryotes and eucaryotes, to produce large quantities of active synthetic proteins comprising truncated analogs, muteins, fusion proteins, and other constructs mimicking the biological activity of the native forms and capable of inducing bone formation in mammals including humans.

The synthetic proteins are useful in clinical applications in conjunction with a suitable delivery or support system (matrix). The matrix is made up of particles or porous materials. The pores must be of a dimension to permit progenitor cell migration and subsequent differentiation and proliferation. The particle size should be within the range of 70 - 850 μm , preferably 70 - 420 μm . It may be fabricated by close packing particulate material into a shape spanning the bone defect, or by otherwise structuring as desired a material that is biocompatible (non-inflammatory) and, biodegradable in vivo to serve as a "temporary scaffold" and substratum for recruitment of migratory progenitor cells, and as a base for their subsequent anchoring and proliferation. Currently preferred carriers include particulate, demineralized, guanidine extracted, species-specific (allogenic) bone, and particulate, deglycosylated (or HF treated), protein extracted, demineralized, xenogenic bone.

Optionally, such xenogenic bone powder matrices also may be treated with proteases such as trypsin. Other useful matrix materials comprise collagen, homopolymers and copolymers of glycolic acid and lactic acid, hydroxyapatite, tricalcium phosphate and other calcium phosphates.

The osteogenic proteins and implantable osteogenic devices enabled and disclosed herein will permit the physician to obtain optimal predictable bone formation to correct, for example, acquired and congenital craniofacial and other skeletal or dental anomalies (Glowacki et al. (1981) Lancet 1:959-963). The devices may be used to induce local endochondral bone formation in non-union fractures, and in other clinical applications including periodontal applications where bone formation is required. Another potential clinical application is in cartilage repair, for example, in the treatment of osteoarthritis.

- 16 -

Brief Description of the Drawing

The foregoing and other objects of this invention, the various features thereof, as well as the invention itself, may be more fully understood from the following description, when read together with the accompanying drawings, in which:

FIGURE 1 is a comparison of the amino acid sequence of various osteogenic proteins to those of the TGF-beta family. COP1, COP3, COP4, COP5, and COP7 are a family of analogs of synthetic osteogenic proteins developed from the consensus gene that was joined to a leader protein via a hinge region having the sequence D-P-N-G that permitted chemical cleavage at the D-P site (by acid) or N-G (by hydroxylamine) resulting in the release of the analog protein; VGI is a Xenopus protein, DPP is a Drosophila protein; OP1 is a native osteogenic protein; CBMP2a and 2b, and CBMP3 are subparts of proteins disclosed in PCT application 087/01537; MIS is Mullerian inhibitory substance; and "consensus choices" represent various substitutions of amino acids that may be made at various positions in osteogenic proteins;

FIGURE 2A is an E. coli expression vector containing a gene of an osteogenic protein fused to a leader protein;

FIGURE 2B is the DNA sequence comprising a modified trp-LE leader, two Fb domains of protein A, an ASP-PRO cleavage site, and the COP5 sequence;

FIGURES 3A and 3B are photomicrographs of implants showing the histology (day 12) of COP5 active recombinant protein. A is a control (rat matrix alone, 25 mg). B is rat matrix plus COP5, showing +++ cartilage formation and ++ bone formation (see key infra). Similar results are achieved with COP7; and

FIGURE 4 is a schematic representation of the DNA sequence and corresponding amino acid sequence of a consensus gene for osteogenic protein (COP0).

Description

Purification protocols have been developed which enable isolation of the osteogenic protein present in crude protein extracts from mammalian bone. The isolation procedure enables the production of significant quantities of substantially pure osteogenic protein from any mammalian species, provided sufficient amounts of fresh bone from the species is available. The empirical development of the procedure, coupled with the availability of fresh calf bone, enabled isolation of substantially pure bovine osteogenic protein (BOP). BOP has been characterized significantly; its ability to induce cartilage and ultimately endochondral bone growth in cat, rabbit, and rat have been studied; it has been shown to be able to induce the full developmental cascade of bone formation previously ascribed to unknown protein or proteins in heterogeneous bone extracts; and it may be used to induce formation of endochondral bone in orthopedic defects including non-union fractures. In its native form it is a glycosylated, dimeric protein. However, it is active in deglycosylated form. It has been partially sequenced.

Elucidation of the amino acid sequence of BOP enabled the construction of a consensus nucleic acid sequence designed as disclosed herein based on the sequence data, inferred codons for the sequences, and observation of partial homology with known genes.

These consensus sequences have been refined by comparison with the sequences present in certain regulatory genes from drosophila, xenopus, and human followed by point mutation, expression, and assay for activity. This approach has been successful in producing several active totally synthetic constructs not found in nature (as far as applicants are aware) which have true osteogenic activity.

These discoveries enable the construction of DNAs encoding totally novel, non-native protein constructs which individually and combined are capable of producing true endochondral bone. The DNAs may be expressed using well established recombinant DNA technologies in procaryotic or eucaryotic host cells, and the expressed proteins may be oxidized and refolded in vitro if necessary for biological activity.

The design and production of such biosynthetic proteins, the nature of the matrix, and other material aspects concerning the nature, utility, how to make, and how to use the subject matter claimed herein will be further understood from the following, which constitutes the best method currently known for practicing the various aspects of the invention.

CONSENSUS SEQUENCE DESIGN

A synthetic consensus gene shown in FIGURE 4 was designed to encode a consensus protein based on amino acid predictions from homology with the

TGF-beta gene family. The designed consensus sequence was then constructed using known techniques involving assembly of oligonucleotides manufactured in a DNA synthesizer.

Tryptic peptides derived from Bovine Osteogenic Protein isolated by applicants and sequenced by Edman degradation provided amino acid sequences that showed strong homology with the Drosophila DPP protein sequence (as inferred from the gene), the Xenopus VGl protein, and somewhat less homology to inhibin and TGF-beta, as demonstrated below in TABLE 1.

- 21 -

TABLE 1

<u>protein</u>	<u>amino acid sequence</u>	<u>homology</u>
(<u>BOP</u>)	SFDAYYCSGACQFPS ***** * * *	(9/15 matches)
(<u>DPP</u>)	GYDAYYCHGKCPFFL	
<hr/>		
(<u>BOP</u>)	SFDAYYCSGACQFPS * * * * *	(6/15 matches)
(<u>Vgl</u>)	GYMANYCYGECPYPL	
<hr/>		
(<u>BOP</u>)	SFDAYYCSGACQFPS * * * * *	(5/15 matches)
(<u>inhibin</u>)	GYHANYCEGECPSHI	
<hr/>		
(<u>BOP</u>)	SFDAYYCSGACQFPS * * * * *	(4/15 matches)
(<u>TGF-beta</u>)	GYHANFCLGPCPYIW	
<hr/>		
(<u>BOP</u>)	K/RACCVPTELSAISMLYLDEN ***** * * * *	(12/20 matches)
(<u>Vgl</u>)	LPCCVPTKMSPISMLFYDNN	
<hr/>		
(<u>BOP</u>)	K/RACCVPTELSAISMLYLDEN * * * * *	(12/20 matches)
(<u>inhibin</u>)	KSCCVPTKLRPMSMLYYDDG	
<hr/>		
(<u>BOP</u>)	K/RACCVPTELSAISMLYLDE **** * *	(6/19 matches)
(<u>TGF-beta</u>)	APCCVPQALEPLPIVYYVG	
<hr/>		
(<u>BOP</u>)	K/RACCVPTELSAISMLYLDEN ***** * * * *	(12/20 matches)
(<u>DPP</u>)	KACCVPTQLDSVAMLYLNDQ	

- 22 -

(<u>BOP</u>)	LYVDF	
	*****	(5/5 matches)
(<u>DPP</u>)	LYVDF	
<hr/>		
(<u>BOP</u>)	LYVDF	
	*** *	(4/5 matches)
(<u>Vgl</u>)	LYVEF	
<hr/>		
(<u>BOP</u>)	LYVDF	
	** **	(4/5 matches)
(<u>TGF-beta</u>)	LYIDF	
<hr/>		
(<u>BOP</u>)	LYVDF	
	* *	(2/5 matches)
(<u>inhibin</u>)	FFVSF	
<hr/>		

*-match

In determining an appropriate amino acid sequence of an osteogenic protein (from which the nucleic acid sequence can be determined), the following points were considered: (1) the amino acid sequence determined by Edman degradation of natural source osteogenic protein tryptic fragments is ranked highest as long as it has a strong signal and shows homology or conservative changes when aligned with the other members of the gene family; (2) where the sequence matches for all four proteins, it is used in the synthetic gene sequence; (3) matching amino acids in DPP and Vgl are used; (4) If Vgl or DPP diverged but either one were matched by inhibin or by TGF-beta, this matched amino acid is chosen; (5) where all sequences diverged, the DPP sequence is initially chosen, with a later plan of creating the

Vgl sequence by mutagenesis kept as a possibility. In addition, the consensus sequence is designed to preserve the disulfide crosslinking and the apparent structural homology among the related proteins.

RECOMBINANT OSTEOGENIC PROTEIN CONSTRUCTS

This approach resulted in the production of novel recombinant proteins capable of inducing formation of cartilage and endochondral bone comprising a protein structure analogous to or duplicative of the functional domain of the naturally sourced material. The amino acid sequences encoded by the consensus DNA sequences were derived from a family of natural proteins implicated in tissue development. These gene products/proteins are known to exist in active form as dimers and are, in general, processed from a precursor protein to produce an active C-terminal domain of the precursor.

The recombinant osteogenic/chondrogenic proteins are "novel" in the sense that, as far as applicants are aware, they do not exist in nature or, if they do exist, have never before been associated with bone or cartilage formation. The approach to design of these proteins is to employ amino acid sequences, found in the native OP isolates, in polypeptide structures are patterned after certain proteins reported in the literature, or the amino acid sequences inferred from DNAs reported in the literature. Thus, using the design criteria set

forth above, and refining the amino acid sequence as more protein sequence information was learned, a series of synthetic proteins were designed with the hope and intent that they might have osteogenic or chondrogenic activity when tested in the bioassay system disclosed below.

It was noted, for example, that DPP from *Drosophila*, VGl from *Xenopus*, the TGF beta family of proteins, and to a lesser extent, alpha and beta inhibins, had significant homologies with certain of the sequences derived from the naturally sourced OP product. (FIGURE 1.) Study of these proteins led to the realization that a portion of the sequence of each had a structural similarity observable by analysis of the positional relationship of cysteines and other amino acids which have an important influence on three dimensional protein conformation. It was noted that a region of these sequences had a series of seven cysteines, placed very nearly in the same relative positions, and certain other amino acids in sequence as set forth below:

10	20	30	40	50
CXXXXLXVFXDXGWXXWXXXPXGXXAXYCXGXCXXPXXXXXXXXNHAXX				
60	70	80	90	100
QXXVXXXNXXXXPXXCCPXXXXXXXXLXXXXXXXXVXLXXYXXMXVXXCXCX				

wherein each X independently represents an amino acid. Expression experiments of two of these constructs demonstrate activity. Expression experiments with constructs patterned after this template amino acid sequence with a shorter sequence having only six cysteines also show activity:

- 25 -

```

      10      20      30      40      50
    LXVXFDXGWXXWXXXPXGXXAXYCXGXCXXPXXXXXXXXNHAXX
      60      70      80      90     100
    QXXVXXXNXXXXPXXCCXPXXXXXXXXXLXXXXXXXXVXLXXYXXMXVXXCXCX

```

wherein each X independently represents an amino acid. Within these generic structures are a multiplicity of specific sequences which have osteogenic or chondrogenic activity. Preferred structures are those having the amino acid sequence:

```

      10      20      30      40      50
    CKRHPLYVDVFRDVGWNDWIVAPPGYHAFYCHGECPPFLADHLNSTNHAIV
      RKRS K S S L  QE VIS E FD Y  E A AY MPESMKAS  VI
      KE F E K I  DN   L   N S  Q  ITK F P   TL
      Q   A   S   K
      60      70      80      90     100
    QTLVNSVNP GKIPKACCVPTELSAISMLYLDENENVVLKNYQDMVVEGCGCR
      SI HAI SEQV EP  A EQMNSLAI FFNDQDK I RK EE T DA H H
      RF   T   S   K DPV V  Y N S   H RN  RS
      N   S               K   P   E

```

wherein, in each position where more than one amino acid is shown, any one of the amino acids shown may be used. Novel active proteins also are defined by amino acid sequences comprising an active domain beginning at residue number 6 of this sequence, i.e., omitting the N terminal CXXXX, or omitting any of the preferred specific combinations such as CKRHP, CRRKQ, CKRHE, etc, resulting in a construct having only six cysteine residues. After this work, PCT 87/01537 was published, and it was observed that the proteins there identified as BMPII a and b and BMPIII each included a region embodying this generic structure. These proteins were not demonstrated to be osteogenic in the published application. However, applicants

discovered that a subpart of the amino acid sequence of these proteins, properly folded, and implanted as set forth herein, is active. These are disclosed herein as CBMPIIa, CBMPIIb, and CBMPIII. Also, applicants retrieved a previously unreported gene by probing a human genomic DNA library with COPO. This protein was designated OPl. It comprises a region exhibiting the same generic structure.

Thus, the preferred osteogenic proteins are expressed from recombinant DNA and comprise amino acid sequences including any of the following sequences:

Vgl	1	10	20	30	40
	CKKRHLVVEFK-DVGWQNWVIAPOGYMANCYGECPYPLTE				
		50	60	70	
	ILNGSN--H-AILQTLVHSIEPED-IPLPCCVPTKMSP				
	80	90	100		
	ISMLFYDNDNDNVVLRHYENMAVDECGCR				
DPP	1	10	20	30	40
	CRRHSLYVDFS-DVGWDDWIVAPLGDAYYCHGKCPFPLAD				
		50	60	70	
	HFNSTN--H-AVVQTLVNNNNPGK-VPKACCVPTQLDS				
	80	90	100		
	VAMLYLNDQSTVVLKNYQEMTVVGCGCR				
OPl	1	10	20	30	40
	LYVSFR-DLGWQDWIIAPEGYAAYYCEGECAFLNS				
		50	60	70	
	YMNATN--H-AIVQTLVHFINPET-VPKPCCAPTQLNA				
	80	90	100		
	ISVLYFDDSSNVILKKYRNMVVRACGCH				

-5
HQRQA

OP1 1 10 20 30 40
CKKHELYVVSFR-DLGWQDWIIAPEGYAAYYCEGECAFP LNS
 50 60 70
YMNATN--H-AIVQTLVHFINPET-VPKPCCAPTQLNA
 80 90 100
ISVLYFDDSSNVILKKYRNMVVRACGCH

CBMP-2a 1 10 20 30 40
CKRHPLYVDFS-DVGWNDWIVAPPGYHAFYCHGECPPFLAD
 50 60 70
HLNSTN--H-AIVQTLVNSVNS-K-IPKACCVPTLSA
 80 90 100
ISMLYLDENEKVVLKKNYQDMVVEGCGCR

CBMP-2b 1 10 20 30 40
CRRHSLYVDFS-DVGWNDWIVAPPGYQAFYCHGDCPPFLAD
 50 60 70
HLNSTN--H-AIVQTLVNSVNS-S-IPKACCVPTLSA
 80 90 100
ISMLYLDEYDKVVLKKNYQEMVVEGCGCR

CBMP-3 1 10 20 30 40
CARRYLKVDFA-DIGWSEWIIISPKSFDAYYCSGACQFPMPS
 50 60 70
SLKPSN--H-ATIQSIVRAVGVPGIPEPCCVPEKMSS
 80 90 100
LSILFFDENKNVVLKVYPNMTVESCACR

COP1 1 10 20 30 40
LYVDFQRDVGWDDWIIAPVDFDAYYCSGACQFPSAD
 50 60 70
HFNSTN--H-AVVQTLVNNMNPBK-VPKPCCVPTLSA
 80 90 100
ISMLYLDENSTVVLKKNYQEMTVVGCGR

COP3 1 10 20 30 40
LYVDFQRDVGWDDWIVAPPGYQAFYCSGACQFPSAD
 50 60 70
HFNSTN--H-AVVQTLVNNMNPBK-VPKPCCVPTLSA
 80 90 100
ISMLYLDENEKVVLKKNYQEMVVEGCGCR

```

COP4      1      10      20      30      40
           LYVDFS-DVGWDDWIVAPPGYQAFYCSGACQFPSAD
                    50      60      70
           HFNSTN--H-AVVQTLVNNMNPBK-VPKPCCVPTLSA
                    80      90     100
           ISMLYLDENEKVVVLKKNYQEMVVEGCGCR

COP5      1      10      20      30      40
           LYVDFS-DVGWDDWIVAPPGYQAFYCHGECPPFLAD
                    50      60      70
           HFNSTN--H-AVVQTLVNSVNSKI--PKACCVPTLSA
                    80      90     100
           ISMLYLDENEKVVVLKKNYQEMVVEGCGCR

COP7      1      10      20      30      40
           LYVDFS-DVGWDDWIVAPPGYHAFYCHGECPPFLAD
                    50      60      70
           HLNSTN--H-AVVQTLVNSVNSKI--PKACCVPTLSA
                    80      90     100
           ISMLYLDENEKVVVLKKNYQEMVVEGCGCR

COP16     1      10      20      30      40
           CRRHSLYVDFS-DVGWDDWIVAPPGYQAFYCHGECPPFLAD
                    50      60      70
           HFNSTN--H-AVVQTLVNSVNSKI--PKACCVPTLSA
                    80      90     100
           ISMLYLDENEKVVVLKKNYQEMVVEGCGCR
           PKHHSQRARKKNKN

```

As shown in FIGURE 1, these sequences have considerable homology with the alpha and beta inhibins, three forms of TGF beta, and MIS.

Gene Preparation

The synthetic genes designed to express the proteins as described above preferably are produced by assembly of chemically synthesized oligonucleotides. 15-100mer oligonucleotides may be synthesized on a Biosearch DNA Model 8600 Synthesizer, and purified by polyacrylamide gel electrophoresis (PAGE) in Tris-Borate-EDTA buffer (TBE). The DNA is then electroeluted from the gel. Overlapping oligomers may be phosphorylated by T4 polynucleotide kinase and ligated into larger blocks which may also be purified by PAGE. Natural gene sequences and cDNAs also may be used for expression.

Expression

The genes can be expressed in appropriate prokaryotic hosts such as various strains of E. coli and also in bacillus, yeasts, and various animal cells such as CHO, myeloma, etc. Generally, expression may be achieved using many cell types and expression systems well known to those skilled in the art. If the gene is to be expressed in E. coli, it must first be cloned into an expression vector. An expression vector (FIGURE 2A) based on pBR322 and containing a synthetic trp promoter operator and the modified trp LE leader can be opened at the EcoRI and PSTI restriction sites, and a FB-FB COP1, COP3, COP5, and COP7 gene fragments (FIGURE 2B) can be inserted between these sites, where FB is fragment B of Staphylococcal Protein A. The expressed fusion

protein results from attachment of the COP gene to a fragment encoding FB. The COP protein is joined to the leader protein via a hinge region having the sequence Asp-Pro-Asn-Gly. This hinge permits chemical cleavage of the fusion protein with dilute acid at the asp-pro site or cleavage at Asn-Gly with hydroxylamine, resulting in release of the COP protein. For COP16 and OPl, the proteins are expressed as fusion products, using the modified trp-LE as a leader.

Production of Active Proteins

The following procedure was followed for production of active recombinant proteins. E. coli cells containing the fusion proteins were lysed. The fusion proteins were purified by differential solubilization. In the case of the COP1, 3, 4, 5, and 7 fusion proteins, cleavage was with dilute acid, and the resulting cleavage products were passed through a Sephacryl-200HR column. The Sephacryl column separated most of the uncleaved fusion products from the COP1, 3, 4, 5, and 7 analogs. In the case of the COP16 or OPl fusion protein, cleavage was with a more concentrated acid, and an SP-Trisacryl column was used as an additional purification step. The COP or OP fractions were then subjected to HPLC on a semi-prep C-18 column.

Initial conditions for refolding of COP analogs or OPl were at pH 8.0 using Tris, Gu-HCl, dithiothreitol. Final conditions for refolding of

COP analogs were at pH 8.0 using Tris, oxidized glutathione, and lower amounts of Gu-HCl and dithiothreitol. Alternatively, the COP or OP1 proteins are suspended in 50 mM HCl, 6 M guanidine-HCl, pH 8.0, for 18 hours at 4°C. Refolding may not be required if the proteins are expressed in animal cells.

Production of Antisera

Antisera to COP7 and COP5 were produced in New Zealand white rabbits. Western blots demonstrate that the antisera react with COP7 and COP5 preparations. Antisera to COP7 has been tested for reactivity to naturally sourced bovine osteogenic protein samples. Western blots show a clear reaction with the 30 kD protein and, when reduced, with the 16 kD subunit. The immunoreactive species appears as a closely-spaced doublet in the 16 kD subunit region, similar to the 16 kD doublet seen in Con A blots.

MATRIX PREPARATION

General Consideration of Matrix Properties

The carrier described in the bioassay section, *infra*, may be replaced by either a biodegradable-synthetic or synthetic-inorganic matrix (e.g., HAP, collagen, tricalcium phosphate, or polylactic acid, polyglycolic acid and various copolymers thereof). Also xenogeneic bone may be used if pretreated as described below.

Studies have shown that surface charge, particle size, the presence of mineral, and the methodology for combining matrix and osteogenic protein all play a role in achieving successful bone induction. Perturbation of the charge by chemical modification abolishes the inductive response. Particle size influences the quantitative response of new bone; particles between 75 and 420 μm elicit the maximum response. Contamination of the matrix with bone mineral will inhibit bone formation. Most importantly, the procedures used to formulate osteogenic protein onto the matrix are extremely sensitive to the physical and chemical state of both the osteogenic protein and the matrix.

The sequential cellular reactions at the interface of the bone matrix/OP implants are complex. The multistep cascade includes: binding of fibrin and fibronectin to implanted matrix, chemotaxis of cells, proliferation of fibroblasts, differentiation into chondroblasts, cartilage formation, vascular invasion, bone formation, remodeling, and bone marrow differentiation.

A successful carrier for osteogenic protein must perform several important functions. It must bind osteogenic protein and act as a slow release delivery system, accommodate each step of the cellular response during bone development, and protect the osteogenic protein from nonspecific proteolysis. In addition, selected materials must be

biocompatible in vivo and biodegradable; the carrier must act as a temporary scaffold until replaced completely by new bone. Biocompatibility requires that the matrix not induce significant inflammation when implanted and not be rejected by the host animal. Biodegradability requires that the matrix be slowly absorbed by the body of the host during development of new bone or cartilage. Polylactic acid (PLA), polyglycolic acid (PGA), and various combinations have different dissolution rates in vivo. In bones, the dissolution rates can vary according to whether the implant is placed in cortical or trabecular bone.

Matrix geometry, particle size, the presence of surface charge, and porosity or the presence of interstices among the particles of a size sufficient to permit cell infiltration, are all important to successful matrix performance. It is preferred to shape the matrix to the desired form of the new bone and to have dimensions which span non-union defects. Rat studies show that the new bone is formed essentially having the dimensions of the device implanted.

The matrix may comprise a shape-retaining solid made of loosely adhered particulate material, e.g., with collagen. It may also comprise a molded, porous solid, or simply an aggregation of close-packed particles held in place by surrounding tissue. Masticated muscle or other tissue may also be used. Large allogeneic bone implants can act as a

carrier for the matrix if their marrow cavities are cleaned and packed with particles and the dispersed osteogenic protein.

Preparation of Biologically Active Allogenic Matrix

Demineralized bone matrix is prepared from the dehydrated diaphyseal shafts of rat femur and tibia as described herein to produce a bone particle size which pass through a 420 μm sieve. The bone particles are subjected to dissociative extraction with 4 M guanidine-HCl. Such treatment results in a complete loss of the inherent ability of the bone matrix to induce endochondral bone differentiation. The remaining insoluble material is used to fabricate the matrix. The material is mostly collagenous in nature, and upon implantation, does not induce cartilage and bone. All new preparations are tested for mineral content and false positives before use. The total loss of biological activity of bone matrix is restored when an active osteoinductive protein fraction or a pure protein is reconstituted with the biologically inactive insoluble collagenous matrix. The osteoinductive protein can be obtained from any vertebrate, e.g., bovine, porcine, monkey, or human, or produced using recombinant DNA techniques.

Preparation of Deglycosylated Bone Matrix for Use in
Xenogenic Implant

When osteogenic protein is reconstituted with collagenous bone matrix from other species and implanted in rat, no bone is formed. This suggests that while the osteogenic protein is xenogenic (not species specific), while the matrix is species specific and cannot be implanted cross species perhaps due to intrinsic immunogenic or inhibitory components. Thus, heretofore, for bone-based matrices, in order for the osteogenic protein to exhibit its full bone inducing activity, a species specific collagenous bone matrix was required.

The major component of all bone matrices is Type I collagen. In addition to collagen, extracted bone includes non-collagenous proteins which may account for 5% of its mass. Many non-collagenous components of bone matrix are glycoproteins. Although the biological significance of the glycoproteins in bone formation is not known, they may present themselves as potent antigens by virtue of their carbohydrate content and may constitute immunogenic and/or inhibitory components that are present in xenogenic matrix.

It has now been discovered that a collagenous bone matrix may be used as a carrier to effect bone inducing activity in xenogenic implants, if one first removes the immunogenic and inhibitory

components from the matrix. The matrix is deglycosylated chemically using, for example, hydrogen fluoride to achieve this purpose.

Bovine bone residue prepared as described above is sieved, and particles of the 74-420 μ M are collected. The sample is dried in vacuo over P_{205} , transferred to the reaction vessel and anhydrous hydrogen fluoride (HF) (10-20 ml/g of matrix) is then distilled onto the sample at -70°C . The vessel is allowed to warm to 0°C . and the reaction mixture is stirred at this temperature for 120 min. After evaporation of the HF in vacuo, the residue is dried thoroughly in vacuo over KOH pellets to remove any remaining traces of acid.

Extent of deglycosylation can be determined from carbohydrate analysis of matrix samples taken before and after treatment with HF, after washing the samples appropriately to remove non-covalently bound carbohydrates.

The deglycosylated bone matrix is next treated as set forth below:

- 1) suspend in TBS (Tris-buffered Saline) 1g/200 ml and stir at 4°C for 2 hrs; or in 6 M urea, 50 mM Tris-HCl, 500 mM NaCl, pH 7.0 (UTBS) and stir at RT for 30 min;

- 2) centrifuge and wash with TBS or UTBS as in step 1; and
- 3) centrifuge; discard supernatant; water wash residue; and then lyophilize.

FABRICATION OF DEVICE

Fabrication of osteogenic devices using any of the matrices set forth above with any of the osteogenic proteins described above may be performed as follows.

A. Ethanol precipitation

In this procedure, matrix is added to osteogenic protein in guanidine-HCl. Samples are vortexed and incubated at a low temperature. Samples are then further vortexed. Cold absolute ethanol is added to the mixture which is then stirred and incubated. After centrifugation (microfuge high speed) the supernatant is discarded. The reconstituted matrix is washed with cold concentrated ethanol in water and then lyophilized.

B. Acetonitrile Trifluoroacetic Acid Lyophilization

In this procedure, osteogenic protein in an acetonitrile trifluoroacetic acid (ACN/TFA) solution is added to the carrier. Samples are vigorously vortexed many times and then lyophilized.

C. Urea Lyophilization

For those proteins that are prepared in urea buffer, the protein is mixed with the matrix, vortexed many times, and then lyophilized. The lyophilized material may be used "as is" for implants.

IN VIVO RAT BIOASSAY

Several of the synthetic proteins have been incorporated in matrices to produce osteogenic devices, and assayed in rat for endochondral bone. Studies in rats show the osteogenic effect to be dependent on the dose of osteogenic protein dispersed in the osteogenic device. No activity is observed if the matrix is implanted alone. The following sets forth guidelines for how the osteogenic devices disclosed herein can be assayed for evaluating protein constructs and matrices for biological activity.

A. Subcutaneous Implantation

The bioassay for bone induction as described by Sampath and Reddi (Proc. Natl. Acad. Sci. USA (1983) 80: 6591-6595), herein incorporated by reference, is used to assess endochondral bone differentiation activity. This assay consists of implanting the test samples in subcutaneous sites in

allogenic recipient rats under ether anesthesia. Male Long-Evans rats, aged 28-32 days, were used. A vertical incision (1 cm) is made under sterile conditions in the skin over the thoracic region, and a pocket is prepared by blunt dissection. Approximately 25 mg of the test sample is implanted deep into the pocket and the incision is closed with a metallic skin clip. The day of implantation is designated as day one of the experiment. Implants were removed on day 12. The heterotopic site allows for the study of bone induction without the possible ambiguities resulting from the use of orthotopic sites.

B. Cellular Events

The implant model in rats exhibits a controlled progression through the stages of matrix induced endochondral bone development including: (1) transient infiltration by polymorphonuclear leukocytes on day one; (2) mesenchymal cell migration and proliferation on days two and three; (3) chondrocyte appearance on days five and six; (4) cartilage matrix formation on day seven; (5) cartilage calcification on day eight; (6) vascular invasion, appearance of osteoblasts, and formation of new bone on days nine and ten; (7) appearance of osteoblastic and bone remodeling and dissolution of the implanted matrix on days twelve to eighteen; and (8) hematopoietic bone marrow differentiation in the ossicle on day twenty-one. The results show that the shape of the new bone conforms to the shape of the implanted matrix.

C. Histological Evaluation

Histological sectioning and staining is preferred to determine the extent of osteogenesis in the implants. Implants are fixed in Bouins Solution, embedded in parafilm, cut into 6-8 mm sections. Staining with toluidine blue or hemotoxylin/eosin demonstrates clearly the ultimate development of endochondrial bone. Twelve day implants are usually sufficient to determine whether the implants show bone inducing activity.

D. Biological Markers

Alkaline phosphatase activity may be used as a marker for osteogenesis. The enzyme activity may be determined spectrophotometrically after homogenization of the implant. The activity peaks at 9-10 days in vivo and thereafter slowly declines. Implants showing no bone development by histology should have little or no alkaline phosphatase activity under these assay conditions. The assay is useful for quantitation and obtaining an estimate of bone formation very quickly after the implants are removed from the rat. Alternatively the amount of bone formation can be determined by measuring the calcium content of the implant.

The osteogenic activity due to osteogenic protein is represented by "bone forming units". One bone forming unit represents the amount of protein that is needed for half maximal bone forming activity as compared to rat demineralized bone matrix as control and determined by calcium content of the implant on day 12.

Devices that contained only rat carrier show complete absence of new bone formation. The implant consists of carrier rat matrix and surrounding mesenchymal cells. Again, the devices that contained rat carrier and not correctly folded (or biologically inactive) recombinant protein also showed complete absence of bone formation. These implants are scored as cartilage formation (-) and bone formation (-). The endochondral bone formation activity is scored as zero percent (0%) (FIGURE 3A).

Implants included biologically active recombinant protein, however, showed evidence of endochondral bone formation. Histologically they showed new cartilage and bone formation.

The cartilage formation is scored as (+) by the presence of metachromatically stained chondrocytes in the center of the implant, as (++) by the presence of numerous chondrocytes in many areas of the implant and as (+++) by the presence of abundant chondrocytes forming cartilage matrix and the appearance of hypertrophied chondrocytes accompanying cartilage calcification (FIGURE 3B).

The bone formation is scored as (+) by the presence of osteoblast surrounding vascular endothelium forming new matrix, as (++) by the formation of bone due to osteoblasts (as indicated by arrows) and further bone remodeling by the appearance of osteoclasts in opposition to the newly formed bone matrix. Vascular invasion is evident in these implants (FIGURE 3B). Formation is scored as (+++) by the presence of extensive remodeled bone which results in the formation of ossicles.

The overall bone inducing activity due to recombinant protein is represented as percent response of endochondral bone formation (see TABLE 2 below).

TABLE 2

HISTOLOGICAL EVALUATION OF RECOMBINANT BONE INDUCTIVE
PROTEINS

<u>Sample</u> <u>No.</u>	<u>Implanted</u> <u>Protein</u>	<u>Cartilage</u> <u>Formation</u>	<u>Bone</u> <u>Formation</u>
260-54	COP-5	+++	++
279-5	COP-5	++	+
285-13	COP-5	+++	++
277-7	COP-7	+++	++
277-8	COP-7	+++	++
277-9	COP-7	++	+
285-14	COP-7	+++	++
285-24	COP-7	++	+
285-25	COP-7	++	++
314-6	COP-16	+++	+++
314-15	COP-16	++	+
314-16	COP-16	++	+
314-12	OP-1	++	+

The invention may be embodied in other specific forms without departing from the spirit or essential characteristics thereof. The present embodiments are therefore to be considered in all respects as illustrative and not restrictive, the scope of the invention being indicated by the appended claims rather than by the foregoing description, and all

- 44 -

changes which come within the meaning and range of equivalency of the claims are therefore intended to be embraced therein.

Claims

1. An osteogenic device for implantation in a mammal, said device comprising:

a biocompatible, in vivo biodegradable matrix defining pores of a dimension sufficient to permit influx, proliferation and differentiation of migratory progenitor cells from the body of said mammal; and

a protein, produced by expression of recombinant DNA in a host cell, comprising one or more polypeptide chains, each of which comprises an amino acid sequence sufficiently duplicative of the sequence of COP5, COP7, COP16, or OP1 such that said protein is capable of inducing endochondral bone formation in association with said matrix when implanted in a mammal.

2. A device for implantation in a mammal, said device comprising:

a biocompatible, in vivo biodegradable matrix defining pores of a dimension sufficient to permit influx, proliferation and differentiation of migratory progenitor cells from the body of said mammal; and

- 46 -

a protein, produced by expression of recombinant DNA in a host cell, comprising one or more polypeptide chains, each of which has less than about 200 amino acids, and each of which comprises a sequence sufficiently duplicative of the sequence of COP5, COP7, COP16, or OP1 such that said protein is capable of inducing cartilage formation in association with said matrix when implanted in a mammal.

3. The device of claim 1 or 2 wherein the sequence comprises:

```

      10      20      30      40      50
CXXXLXVFXDXGWXXWXXXPXGXXAXYCXGXCXXPXXXXXXXXNHAXX
      60      70      80      90     100
QXXVXXNXXXXPXXCCXPXXXXXXXXLXXXXXXVXLXXYXXMXVXXCXCX

```

wherein each X independently represents an amino acid.

4. The device of claim 1 or 2 wherein the sequence comprises:

```

      10      20      30      40      50
LXVFXDXGWXXWXXXPXGXXAXYCXGXCXXPXXXXXXXXNHAXX
      60      70      80      90     100
QXXVXXNXXXXPXXCCXPXXXXXXXXLXXXXXXVXLXXYXXMXVXXCXCX

```

wherein each X independently represents an amino acid.

5. The device of claim 1 or 2 wherein the sequence comprises:

```

      10      20      30      40      50
CKRHPLYVDFRDVGWNDWIVAPPGYHAFYCHGECPPFLADHLNSTNHAIV
RKRS K S S L QE VIS E FD Y E A AY MPESMKAS VI
KE F E K I DN L N S Q ITK F P TL
Q A S K
      60      70      80      90      100
QTLVNSVNP GKIPKACCVPTELSAISMLYLDENENVVLKNYQDMVVEGCGCR
SI HAI SEQV EP A EQMNSLAI FFNDQDK I RK EE T DA H H
RF T S K DPV V Y N S H RN RS
N S K P E

```

wherein, in each position where more than one amino acid is shown, any one of the amino acids shown may be in that position.

6. The device of claim 1 or 2 wherein the sequence comprises:

```

      10      20      30      40      50
LYVDFRDVGWNDWIVAPPGYHAFYCHGECPPFLADHLNSTNHAIV
K S S L QE VIS E FD Y E A AY MPESMKAS VI
F E K I DN L N S Q ITK F P TL
A S K
      60      70      80      90      100
QTLVNSVNP GKIPKACCVPTELSAISMLYLDENENVVLKNYQDMVVEGCGCR
SI HAI SEQV EP A EQMNSLAI FFNDQDK I RK EE T DA H H
RF T S K DPV V Y N S H RN RS
N S K P E

```

wherein, in each position where more than one amino acid is shown, any one of the amino acids shown may be in that position.

7. The device of claim 1 or 2 wherein the sequence comprises:

```

      1      10      20      30      40
Vg1  CKKRHLYVEFK-DVGWQNWVIAPQGYMANYCYGECYPPLTE
      50      60      70
      ILNGSN--H-AILQTLVHSIEPED-IPLPCCVPTKMSP
      80      90      100
      ISMLFYDNNNDNVVLRHYENMAVDECGCR

```

8. The device of claim 1 or 2 wherein the sequence comprises:

```

      1      10      20      30      40
DPP  CRRHSLYVDFS-DVGWDDWIVAPLGDAYYCHGKCPFLAD
      50      60      70
      HFNSTN--H-AVVQTLVNNNNPGK-VPKACCVPTQLDS
      80      90      100
      VAMLYLNDQSTVVLKNYQEMTVVGCGR

```

9. The device of claim 1 or 2 wherein the sequence comprises:

```

      1      10      20      30      40
OP1  LYVSFR-DLGWQDWIIAPEGYAAYYCEGECAFLNS
      50      60      70
      YMNATN--H-AIVQTLVHFINPET-VPKPCCAPTQLNA
      80      90      100
      ISVLYFDDSSNVILKKYRNMVVRACGCH

```

10. The device of claim 1 or 2 wherein the sequence comprises:

				-5
				HQRQA
OP1	1	10	20	30
				40
				CKKHELYVSFR-DLGWQDWIIAPEGYAAYYCEGECAFLNS
		50	60	70
				YMNATN--H-AIVQTLVHFNPET-VPKPCCAPTQLNA
	80	90	100	
				ISVLYFDDSSNVILKKYRNMVVRACGCH

11. The device of claim 1 or 2 wherein the sequence comprises:

				30	40
CBMP-2a	1	10	20		
					CKRHPLYVDFS-DVGWNDWIVAPPGYHAFYCHGECPPFLAD
		50	60	70	
					HLNSTN--H-AIVQTLVNSVNS-K-IPKACCVPTLSA
	80	90	100		
					ISMLYLDENEKVVLKNYQDMVVEGCGCR

12. The device of claim 1 or 2 wherein the sequence comprises:

				30	40
CBMP-2b	1	10	20		
					CRRHSLYVDFS-DVGWNDWIVAPPGYQAFYCHGDCPPFLAD
		50	60	70	
					HLNSTN--H-AIVQTLVNSVNS-S-IPKACCVPTLSA
	80	90	100		
					ISMLYLDEYDKVVLKNYQEMVVEGCGCR

- 50 -

13. The device of claim 1 or 2 wherein the sequence comprises:

	1	10	20	30	40
CBMP-3	CARRYLKVDFA-DIGWSEWIISPKSFDAYYCSGACQFPMPK				
		50	60	70	
	SLKPSN--H-ATIQSIVRAVGVPGIPEPCCVPEKMSS				
	80	90	100		
	LSILFFDENKNVVLKVYPNMTVESACR				

14. The device of claim 1 or 2 wherein the sequence comprises:

	1	10	20	30	40
COP1	LYVDFQRDVGWDDWIIAPVDFDAYYCSGACQFPSAD				
		50	60	70	
	HFNSTN--H-AVVQTLVNNMNPBK-VPKPCCVPTLSA				
	80	90	100		
	ISMLYLDENSTVVLKNYQEMTVVGCGR				

15. The device of claim 1 or 2 wherein the sequence comprises:

	1	10	20	30	40
COP3	LYVDFQRDVGWDDWIVAPPGYQAFYCSGACQFPSAD				
		50	60	70	
	HFNSTN--H-AVVQTLVNNMNPBK-VPKPCCVPTLSA				
	80	90	100		
	ISMLYLDENEKVVLKNYQEMVVEGCGCR				

16. The device of claim 1 or 2 wherein the sequence comprises:

	1	10	20	30	40
COP4	LYVDFS-DVGWDDWIVAPPGYQAFYCSGACQFPSAD				
		50	60	70	
	HFNSTN--H-AVVQTLVNNMNPBK-VPKPCCVPTLSA				
	80	90	100		
	ISMLYLDENEKVVLKNYQEMVVEGCGCR				

17. The device of claim 1 or 2 wherein the sequence comprises:

```

      1          10          20          30          40
COP5      LYVDFS-DVGWDDWIVAPPGYQAFYCHGECPPFLAD
              50          60          70
      HFNSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA
              80          90          100
      ISMLYLDENEKVVLKNYQEMVVEGCGCR
  
```

18. The device of claim 1 or 2 wherein the sequence comprises:

```

      1          10          20          30          40
COP7      LYVDFS-DVGWNDWIVAPPGYHAFYCHGECPPFLAD
              50          60          70
      HLNSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA
              80          90          100
      ISMLYLDENEKVVLKNYQEMVVEGCGCR
  
```

19. The device of claim 1 or 2 wherein the sequence comprises:

```

                                10
                                PKHHSQRARKKNKN
      1          10          20          30          40
COP16     CRRHSLYVDFS-DVGWNDWIVAPPGYQAFYCHGECPPFLAD
              50          60          70
      HFNSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA
              80          90          100
      ISMLYLDENEKVVLKNYQEMVVEGCGCR
  
```

20. The device of claim 1 or 2 wherein the osteogenics protein comprises a pair of separate polypeptide chains.

- 52 -

21. Osteogenic protein, produced by expression of recombinant DNA in a host cell, comprising one or more polypeptide chains each of which comprises a sequence sufficiently duplicative of the sequence of COP5, COP7, COP16, or OPl such that said protein is capable of inducing endochondral bone formation in association with a matrix when implanted in a mammal.

22. A protein, produced by expression of recombinant DNA in a host cell, comprising one or more polypeptide chains less than about 200 amino acids long, and comprising a sequence sufficiently duplicative of the sequence of COP5, COP7, COP16, or OPl such that said protein is capable of inducing cartilage formation in association with a matrix when implanted in a mammal.

23. The protein of claim 21 or 22 further characterized by being unglycosylated.

24. The protein of claim 21 or 22 comprising a pair of separate polypeptide chains.

25. The protein of claim 21 or 22 comprising the amino acid sequences:

10	20	30	40	50
CXXXXLXVXF	DXGWXXWXXX	PXGXXAXYC	GXCPXXXXXX	NHAXX
60	70	80	90	100
QXXVXXXNXXX	PXXCCXPXXXXX	XXXXLXXXXXX	VXLXXYXXM	VXXCXCX

wherein each X independently represents an amino acid.

26. The protein of claim 21 or 22 comprising the amino acid sequences:

10	20	30	40	50
LXVXF	DXGWXXWXXX	PXGXXAXY	CXGXCXXP	XXXXXXXXNHAXX
60	70	80	90	100
QXXVXXXN	XXXXPXXCC	PXXXXXXX	XXXXXXVXL	XXYXXMXVXXCXCX

wherein each X independently represents an amino acid.

27. The protein of claim 21 or 22 comprising the amino acid sequences:

10	20	30	40	50
CKRHPLYV	DFRDVGW	NDWIVAP	PGYHAFY	CHGECPF
RKRS K S S L	QE VIS E	FD Y E A	AY MPES	MKAS VI
KE F E K I	DN L N S	Q ITK F P	TL	
Q A S K				
60	70	80	90	100
QTLVNSV	NPGKIPK	ACCVPT	ELSAIS	MLYLDEN
SI HAI	SEQV EP	A EQMNS	LAI FFND	QDK I RK
RF T S	K DPV V	Y N S H	RN RS	
N S		K P E		

wherein, in each position where more than one amino acid is shown, any one of the amino acids shown may be in that position.

- 54 -

28. The protein of claim 21 or 22 comprising the amino acid sequences:

```

      10      20      30      40      50
LYVDFRDVGWNDWIVAPPGYHAFYCHGECPPFLADHLNSTNHAIV
  K S S L  QE VIS E FD Y  E A AY MPESMKAS  VI
  F E K I  DN      L   N S   Q  ITK F P   TL
      A      S      K
      60      70      80      90      100
QTLVNSVNP GKIPKACCVPTELSAISMLYLDENENVVLKQDMVVEGCGCR
SI HAI SEQV EP  A EQMNSLAI FFNDQDK I RK EE T DA H H
  RF   T   S      K DPV V  Y N S      H RN  RS
  N   S                               K      P   E

```

wherein, in each position where more than one amino acid is shown, any one of the amino acids shown may be in that position.

29. The protein of claim 21 or 22 comprising the amino acid sequences:

```

      1      10      20      30      40
Vg1  CKKRHL YVEFK-DVGWQNWVIA PQGYMAN YCYGEC PYPLTE
      50      60      70
      ILNGSN--H-AILQTLVHSIEPED-IPLPCCVPTKMSP
      80      90      100
      ISMLFYDNNDNVVLRHYENMAVDECGCR

```

30. The protein of claim 21 or 22 comprising the amino acid sequences:

```

      1      10      20      30      40
DPP  CRRHSL YVDFS-DVGWDDWIVAPLGYDAY YCHGKCPFLAD
      50      60      70
      HFNSTN--H-AVVQTLVNNNNPGK-VPKACCVPTQLDS
      80      90      100
      VAMLYLNDQSTVVLKQYQEMTVVGCGCR

```


31. The protein of claim 21 or 22 comprising the amino acid sequence:

```

      1      10      20      30      40
OP1      LYVSFR-DLGWQDWIIAPEGYAAYYCEGECAPFLNS
              50      60      70
      YMNATN--H-AIVQTLVHFINPET-VPKPCCAPTQLNA
              80      90      100
      ISVLYFDDSSNVILKKYRNMVVRACGCH

```

32. The protein of claim 21 or 22 comprising the amino acid sequences:

```

                                          -5
                                          HQRQA
      1      10      20      30      40
OP1      CKKHELYVSFR-DLGWQDWIIAPEGYAAYYCEGECAPFLNS
              50      60      70
      YMNATN--H-AIVQTLVHFINPET-VPKPCCAPTQLNA
              80      90      100
      ISVLYFDDSSNVILKKYRNMVVRACGCH

```

33. The protein of claim 21 or 22 comprising the amino acid sequences:

```

      1      10      20      30      40
CBMP-2a  CKRHPLYVDFS-DVGWNDWIVAPPGYHAFYCHGECPFPLAD
              50      60      70
      HLNSTN--H-AIVQTLVNSVNS-K-IPKACCVPTLSA
              80      90      100
      ISMLYLDENEKVVVLKKNYQDMVVEGCGCR

```

34. The protein of claim 21 or 22 comprising the amino acid sequences:

```

      1      10      20      30      40
CBMP-2b CRRHSLYVDFS-DVGWNDWIVAPPGYQAFYCHGDCPFPLAD
              50      60      70
      HLNSTN--H-AIVQTLVNSVNS-S-IPKACCVPTELSA
              80      90     100
      ISMLYLDEYDKVVLKQEMVVEGCGCR
  
```

35. The protein of claim 21 or 22 comprising the amino acid sequences:

```

      1      10      20      30      40
CBMP-3  CARRYLKVDFA-DIGWSEWIIISPKSFDAYYCSGACQFPMK
              50      60      70
      SLKPSN--H-ATIQSIVRAVGVPPIPEPCCVPEKMSS
              80      90     100
      LSILFFDENKNVVLKVYPNMTVESCACR
  
```

36. The protein of claim 21 or 22 comprising the amino acid sequences:

```

      1      10      20      30      40
COP1    LYVDFQRDVGWDDWIIAPVDFDAYYCSGACQFPSAD
              50      60      70
      HFNSTN--H-AVVQTLVNNMNPVK-VPKPCCVPTELSA
              80      90     100
      ISMLYLDENSTVVLKQEMTVVVGCGCR
  
```

37. The protein of claim 21 or 22 comprising the amino acid sequences:

```

      1      10      20      30      40
COP3    LYVDFQRDVGWDDWIVAPPGYQAFYCSGACQFPSAD
              50      60      70
      HFNSTN--H-AVVQTLVNNMNPVK-VPKPCCVPTELSA
              80      90     100
      ISMLYLDENEKVVLKQEMVVEGCGCR
  
```

38. The protein of claim 21 or 22 comprising the amino acid sequences:

```

1           10           20           30           40
COP4      LYVDFS-DVGWDDWIVAPPGYQAFYCSGACQFPSAD
           50           60           70
      HFNSTN--H-AVVQTLVNNMNP GK-VPKPCCVPTELSA
           80           90          100
      ISMLYLDENEKVVLKKNYQEMVVEGCGCR

```

39. The protein of claim 21 or 22 comprising the amino acid sequences:

```

1           10           20           30           40
COP5      LYVDFS-DVGWDDWIVAPPGYQAFYCHGECPPFLAD
           50           60           70
      HFNSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA
           80           90          100
      ISMLYLDENEKVVLKKNYQEMVVEGCGCR

```

40. The protein of claim 21 or 22 comprising the amino acid sequences:

```

1           10           20           30           40
COP7      LYVDFS-DVGWDDWIVAPPGYHAFYCHGECPPFLAD
           50           60           70
      HLNSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA
           80           90          100
      ISMLYLDENEKVVLKKNYQEMVVEGCGCR

```

41. The protein of claim 21 or 22 comprising the amino acid sequences:

```

                                           -10
                                           PKHHSSRARKKNKN
1           10           20           30           40
COP16     CRRHSLYVDFS-DVGWDDWIVAPPGYQAFYCHGECPPFLAD
           50           60           70
      HFNSTN--H-AVVQTLVNSVNSKI--PKACCVPTELSA
           80           90          100
      ISMLYLDENEKVVLKKNYQEMVVEGCGCR

```

42. The protein of claim 21 or 22, comprising the product of expression of a DNA in a procaryotic cell.

43. A cell line engineered to express the protein of claim 21 or 22.

44. The protein of claim 21 having a half maximum bone forming activity of about 20 - 25 ng per 25 mg of implant.

45. Antibodies reactive with an epitope of the protein of claim 21 or 22.

46. The device of claim 1 or 2 wherein said matrix comprises demineralized, deglycosylated, protein extracted, particulate, xenogenic bone.

47. The device of claim 1 or 2 wherein said matrix comprises demineralized, protein-extracted, particulate, xenogenic bone treated with HF.

1 / 10

FIG. 1-1									
:::VGL DPP OPI CBMP2 :CBMP3 beta-									
COP1 COP5 COP7 COP4									
COP3									
COP2									
COP1									
COP2									
COP3									
COP4									
COP5									
COP6									
COP7									
COP8									
COP9									
COP10									
COP11									
COP12									
COP13									
COP14									
COP15									
COP16									
COP17									
COP18									
COP19									
COP20									
COP21									
COP22									
COP23									
COP24									
COP25									
COP26									
COP27									
COP28									
COP29									
COP30									
COP31									
COP32									
COP33									
COP34									
COP35									
COP36									
COP37									
COP38									
COP39									
COP40									
COP41									
COP42									
COP43									
COP44									
COP45									
COP46									
COP47									
COP48									
COP49									
COP50									
COP51									
COP52									
COP53									
COP54									
COP55									
COP56									
COP57									
COP58									
COP59									
COP60									
COP61									
COP62									
COP63									
COP64									
COP65									
COP66									
COP67									
COP68									
COP69									
COP70									
COP71									
COP72									
COP73									
COP74									
COP75									
COP76									
COP77									
COP78									
COP79									
COP80									
COP81									
COP82									
COP83									
COP84									
COP85									
COP86									
COP87									
COP88									
COP89									
COP90									
COP91									
COP92									
COP93									
COP94									
COP95									
COP96									
COP97									
COP98									
COP99									
COP100									
COP101									
COP102									
COP103									
COP104									
COP105									
COP106									
COP107									
COP108									
COP109									
COP110									
COP111									
COP112									
COP113									
COP114									
COP115									
COP116									
COP117									
COP118									
COP119									
COP120									

5 / 10

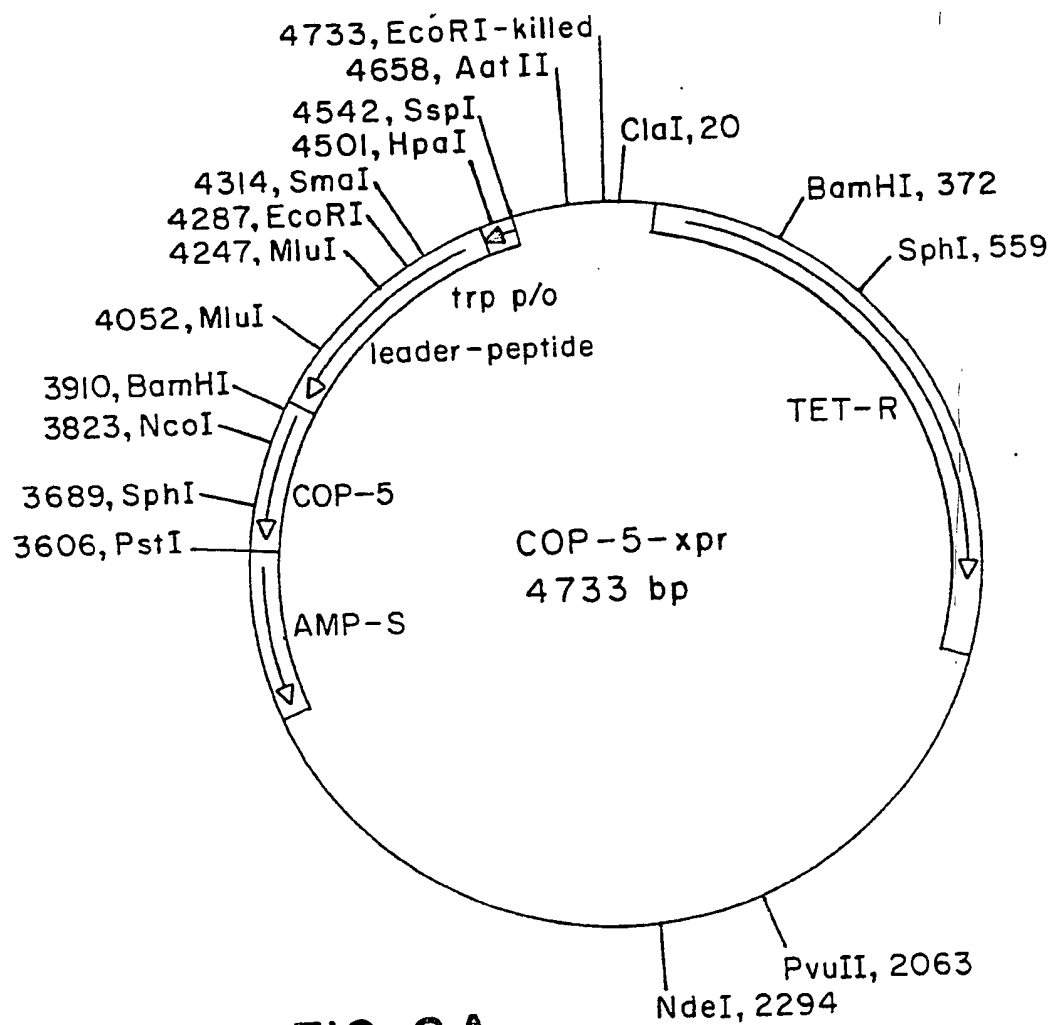


FIG. 2A

6 / 10

FIG. 2B-1

COP-5 fusion protein

10 20 30 40 50
 ATGAAAGCAATTTTCGTACTGAAAGGTTCACTGGACAGAGATCTGGACTC
 M K A I F V L K G S L D R D L D S
 BglII

60 70 80 90 100
 TCGTCTGGATCTGGACGTTTCGTACCGACCACAAAGACCTGTCTGATCACC
 R L D L D V R T D H K D L S D H

110 120 130 140 150
 TGGTTCTGGTCGACCTGGCTCGTAACGACCTGGCTCGTATCGTTACTCCC
 L V L V D L A R N D L A R I V T P
 SalI Sma

160 170 180 190 200
 GGGTCTCGTTACGTTGCGGATCTGGAATTCATGGCTGACAACAAATTCAA
 G S R Y V A D L E F M A D N K F N
 I EcoRI

210 220 230 240 250
 CAAGGAACAGCAGAACGCGTTCTACGAGATCTTGACCTGCCGAACCTGA
 K E Q Q N A F Y E I L H L P N L
 MluI BglII BspMI+

260 270 280 290 300
 ACCGAAGACGCGTAACGGCTTCATCCAAAGCTTGAAGGATGAGCCCTCT
 N E E Q R N G F I Q S L K D E P S
 HindIII

310 320 330 340 350
 CAGTCTGCGAATCTGCTAGCGGATGCCAAGAACTGAACGATGCGCAGGC
 Q S A N L L A D A K K L N D A Q A
 NheI FspI

360 370 380 390 400
 ACCGAAATCGGATCAGGGGCAATTCATGGCTGACAACAAATTCAACAAGG
 P K S D Q G Q F M A D N K F N K

410 420 430 440 450
 AACAGCAGAACGCGTTCTACGAGATCTTGACCTGCCGAACCTGAACGAA
 E Q Q N A F Y E I L H L P N L N E
 MluI BglII BspMI+

460 470 480 490 500
 GAGCAGCGTAACGGCTTCATCCAAAGCTTGAAGGATGAGCCCTCTCAGTC
 E Q R N G F I Q S L K D E P S Q S
 HindIII

SUBSTITUTE SHEET

7 / 10

FIG. 2B-2

510 520 530 540 550
TGC GAATCTGCTAGCGGATGCCAAGAACTGAACGATGCGCAGGCACCGA
A N L L A D A K K L N D A Q A P
NheI FspI

560 570 580 590 600
AGGATCCTAATGGGCTGTACGTCGACTTCAGCGACGTGGGCTGGGACGAC
K D P N G L Y V D F S D V G W D D
BamHI SalI

610 620 630 640 650
TGGATTGTGGCCCCACCAGGCTACCAGGCCTTCTACTGCCATGGCGAATG
W I V A P P G Y Q A F Y C H G E C
StuI NcoI BsmI+

660 670 680 690 700
CCCTTTCCCGCTAGCGGATCACTTCAACAGCACCAACCACGCCGTGGTGC
P F P L A D H F N S T N H A V V
NheI DraIII PflMI

710 720 730 740 750
AGACCCTGGTGAACCTCTGTCAACTCCAAGATCCCTAAGGCTTGCTGCGTG
Q T L V N S V N S K I P K A C C V
MstII

760 770 780 790 800
CCCACCGAGCTGTCCGCCATCAGCATGCTGTACCTGGACGAGAATGAGAA
P T E L S A I S M L Y L D E N E K
SphI

810 820 830 840 850
GGTGGTGTCTGAAGAACTACCAGGAGATGGTAGTAGAGGGCTGCGGCTGCC
V V L K N Y Q E M V V E G C G C
PflMI

860
GCTAACTGCAG
R *
PstI

SUBSTITUTE SHEET

FIG. 3A

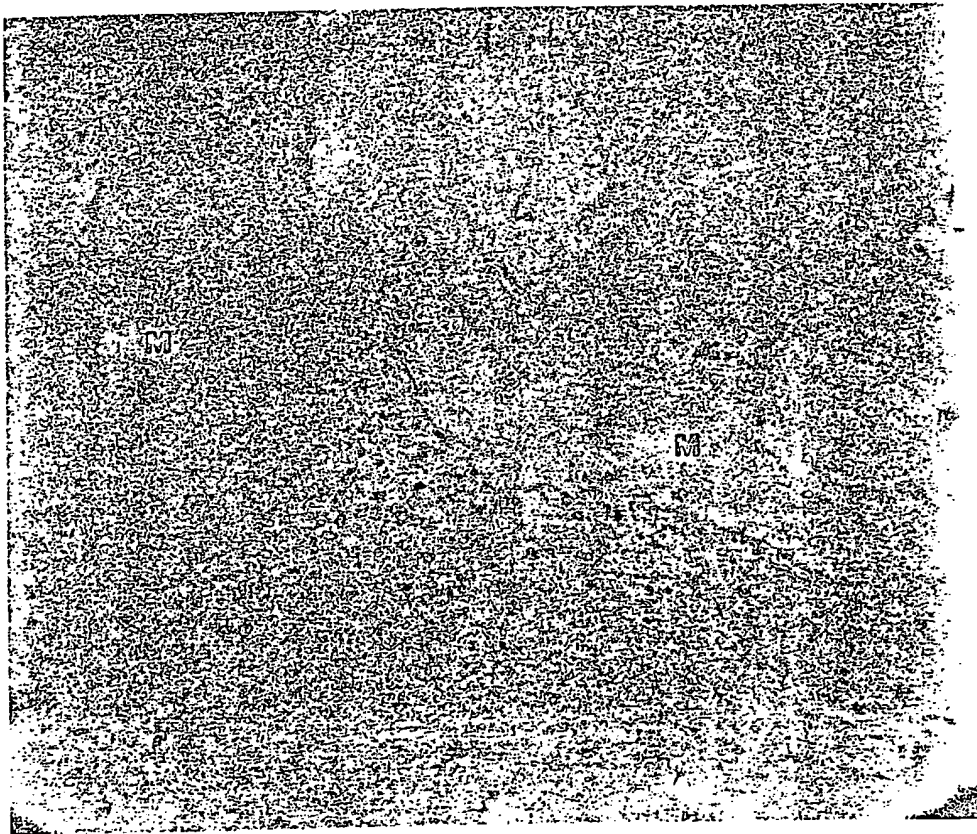


FIG. 3B



10 / 10

FIG. 4

10 20 30 40 50
GATCCTAATGGGCTGTACGTGGACTTCCAGCGCGACGTGGGCTGGGACGA
D P N G L Y V D F Q R D V G W D D

60 70 80 90 100
CTGGATCATCGCCCCCGTCGACTTCGACGCCTACTACTGCTCCGGAGCCT
W I I A P V D F D A Y Y C S G A

110 120 130 140 150
GCCAGTTCCTCTGCGGATCACTTCAACAGCACCAACCACGCCGTGGTG
C Q F P S A D H F N S T N H A V V

160 170 180 190 200
CAGACCCTGGTGAACAACATGAACCCCGGCAAGGTACCCAAGCCCTGCTG
Q T L V N N M N P G K V P K P C C

210 220 230 240 250
CGTGCCCAACCGAGCTGTCCGCCATCAGCATGCTGTACCTGGACGAGAATT
V P T E L S A I S M L Y L D E N

260 270 280 290 300
CCACCGTGGTGCTGAAGAACTACCAGGAGATGACCGTGGTGGGCTGCGGC
S T V V L K N Y Q E M T V V G C G

310
TGCCGCTAACTGCAG
C R *

SUBSTITUTE SHEET